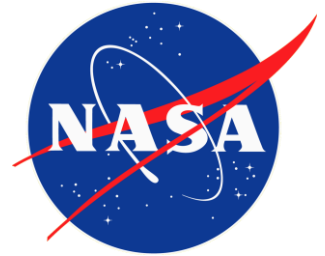


ATMOSPHERIC
SCIENCE
DATA
CENTER



Conversion of Archived HDF Satellite Level 2 Swath Data Products to NetCDF

ESIP 2013 Summer Session
Walt Baskin¹, Jennifer Perez²

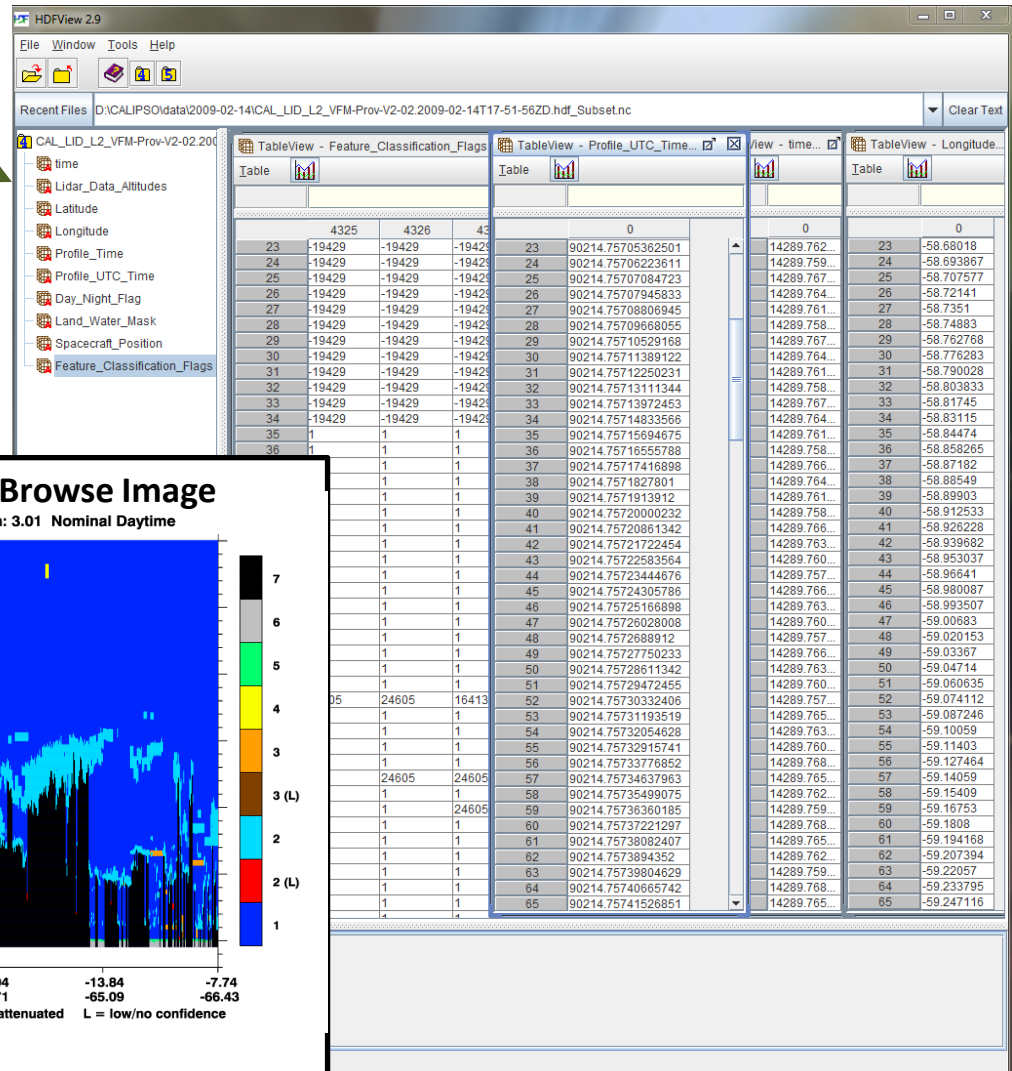
- (1) Science Systems and Applications, Inc. (SSAI), Hampton, VA
- (2) NASA Langley Research Center, Hampton, VA

HDF4 to CF Compliant NetCDF Conversion Challenges

- The NASA Atmospheric Sciences Data Center (ASDC) has developed and deployed provider-specific HDF-to-NetCDF conversion options in search and subset web applications for the CALIPSO, CERES, TES, and MOPITT missions with tailored subsetting for level 2 satellite datasets.
- This presentation explores the challenges in HDF to NetCDF conversion encountered by the ASDC's development team.

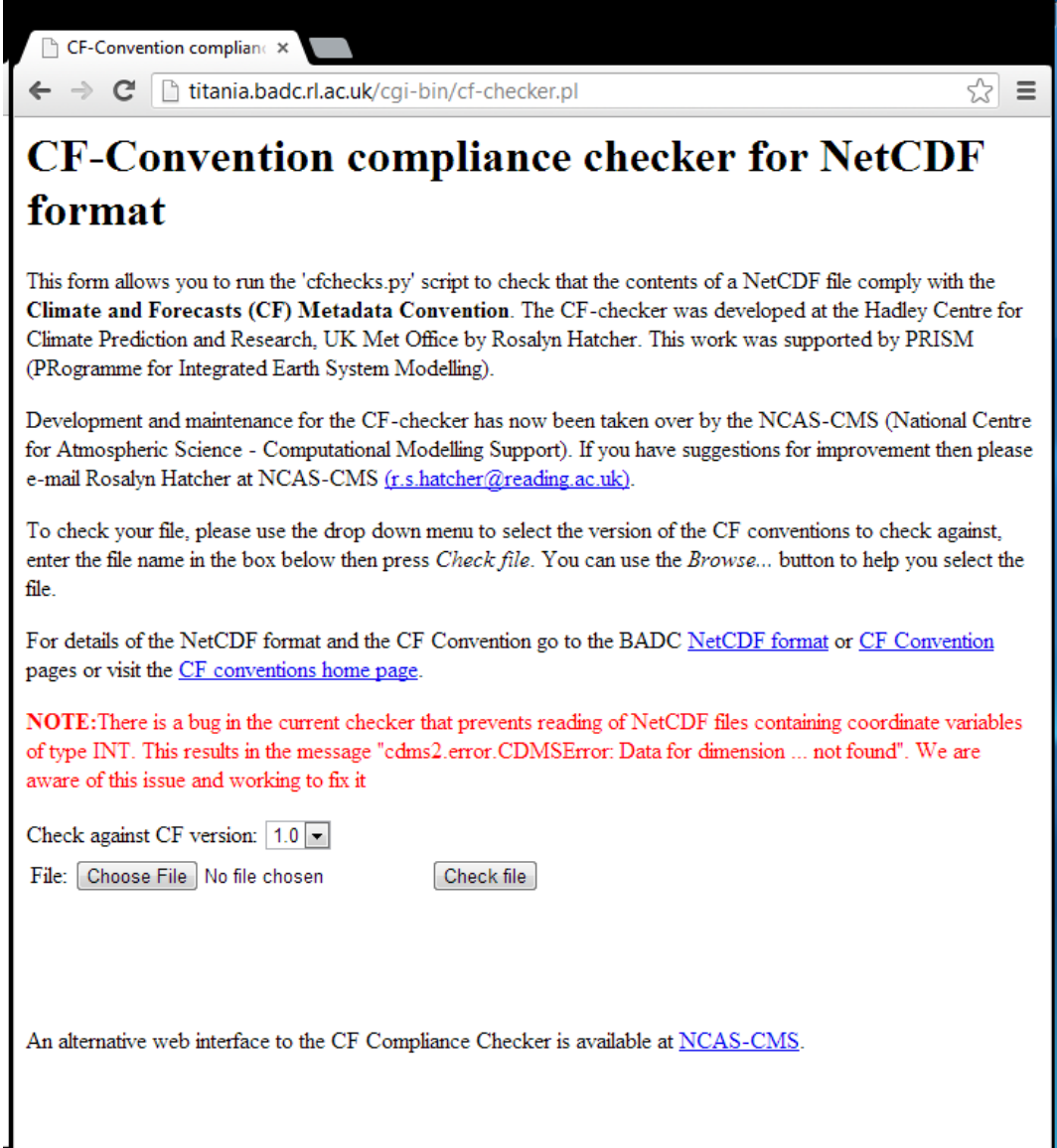
HDF4 to CF Compliant NetCDF Conversion Challenges

ASDC CALIPSO Subsetter NetCDF output of Level2 CALIPSO Vertical Feature Mask (VFM) data granule



HDF4 to CF Compliant NetCDF Conversion Challenges

In order to identify CF compliance problems, the NetCDF result files generated by the ASDC CALIPSO Subsetter were uploaded to the following site running a CF checker developed by Rosalyn Hatcher at the Halley Centre for Climate Prediction and Research.



The screenshot shows a web browser window with the address bar displaying 'titania.badc.rl.ac.uk/cgi-bin/cf-checker.pl'. The page title is 'CF-Convention compliance checker for NetCDF format'. The main content area contains the following text:

This form allows you to run the 'cfchecks.py' script to check that the contents of a NetCDF file comply with the **Climate and Forecasts (CF) Metadata Convention**. The CF-checker was developed at the Hadley Centre for Climate Prediction and Research, UK Met Office by Rosalyn Hatcher. This work was supported by PRISM (PRogramme for Integrated Earth System Modelling).

Development and maintenance for the CF-checker has now been taken over by the NCAS-CMS (National Centre for Atmospheric Science - Computational Modelling Support). If you have suggestions for improvement then please e-mail Rosalyn Hatcher at NCAS-CMS (r.s.hatcher@reading.ac.uk).

To check your file, please use the drop down menu to select the version of the CF conventions to check against, enter the file name in the box below then press *Check file*. You can use the *Browse...* button to help you select the file.

For details of the NetCDF format and the CF Convention go to the BADC [NetCDF format](#) or [CF Convention](#) pages or visit the [CF conventions home page](#).

NOTE: There is a bug in the current checker that prevents reading of NetCDF files containing coordinate variables of type INT. This results in the message "cdms2.error: CDMSException: Data for dimension ... not found". We are aware of this issue and working to fix it

Check against CF version:

File: No file chosen

An alternative web interface to the CF Compliance Checker is available at [NCAS-CMS](#).

HDF4 to CF Compliant NetCDF Conversion Challenges

The screenshot shows the HDFView 2.9 interface. On the left, a tree view lists various data granules, including 'Day_Night_Flag', 'Feature_Classification_Flags', 'Land_Water_Mask', and several 'VDFDim0_vdata_metadata_vdf' and 'VDFDim1_vdata_metadata_vdf' granules. The right pane displays a table view for the 'Feature_Classification_Flag' granule. The table has columns for row numbers and values, showing a mix of 46107 and 1. The bottom pane shows metadata for 'Feature_Classification_Flags (15350, 2)', including dimensions and units.

	4325	4326	4327
23	46107	46107	46107
24	46107	46107	46107
25	46107	46107	46107
26	46107	46107	46107
27	46107	46107	46107
28	46107	46107	46107
29	46107	46107	46107
30	46107	46107	46107
31	46107	46107	46107
32	46107	46107	46107
33	46107	46107	46107
34	46107	46107	46107
35	1	1	1
36	1	1	1
37	1	1	1
38	1	1	1
39	1	1	1
40	1	1	1
41	1	1	1
42	1	1	1
43	1	1	1
44	1	1	1
45	1	1	1
46	1	1	1
47	1	1	1
48	1	1	1
49	1	1	1
50	1	1	1
51	1	1	1
52	24605	24605	16413
53	1	1	1
54	1	1	1
55	1	1	1
56	1	1	1
57	1	24605	24605
58	1	1	1
59	1	1	24605
60	1	1	1
61	1	1	1
62	1	1	1

Feature_Classification_Flags (15350, 2)
16-bit unsigned integer, 225 x 5515
Number of attributes = 6
DIMENSION_LIST = 11158, 12031
format = Uint_1
long_name = Feature_Classification_Flags
origname = Feature_Classification_Flags
units = N/A
valid_range = N/A

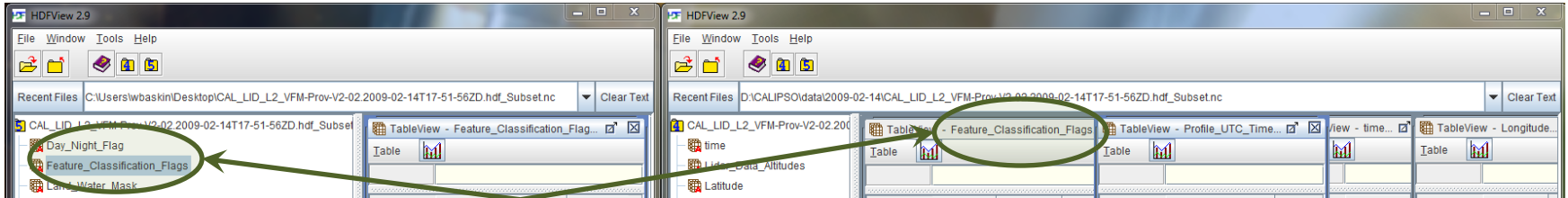
The HDF Group released the HDF4 CF Conversion Toolkit on June 15, 2013.

The HDFView screensots on the left depict NetCDF output from the h410nccf_nc4 utility.

ASDC CALIPSO Subsetter HDF4 output of Level2 CALIPSO Vertical Feature Mask (VFM) data granule

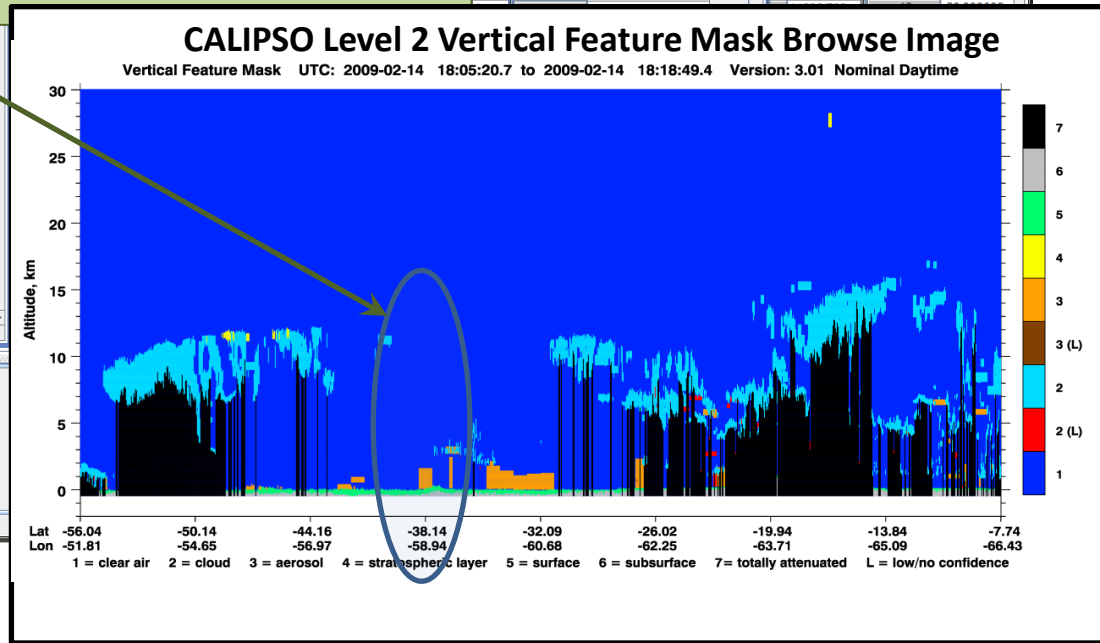
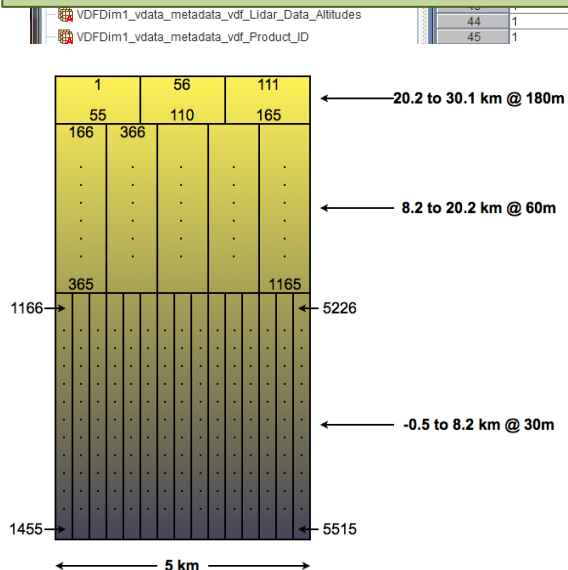
h410nccf_nc4

HDF4 to CF Compliant NetCDF Conversion Challenges



Packed data arrays with varying spatial context

- Visualization and Analysis Clients would not be able to plot the data shown in the Feature Mask Browse image since Feature_Classification_Flags for each vertical profile are packed in a 5515 element array (shown below) and require a non-trivial algorithm to extract the information.



HDF4 to CF Compliant NetCDF Conversion Challenges

No dimension labels in original hdf data granule

- If dimensions are not labeled in the source hdf file a generic name must be provided to be CF compliant. These names are usually used by client applications to label axes of plots.

Feature_Classification_Flags (15350, 2)
16-bit unsigned integer, 225 x 5515
Number of attributes = 6
DIMENSION_LIST = 11158, 12031
format = Uint_1
long_name = Feature_Classification_Flags
orig_name = Feature_Classification_Flags
units = N/A
valid_range = N/A

Feature_Classification_Flags (67687001, 9)
16-bit integer, 225 x 5515
Number of attributes = 3
long_name = Feature_Classification_Flags
units = N/A
format = Uint_1

HDF4 to CF Compliant NetCDF Conversion Challenges

The image displays two side-by-side windows of HDFView 2.9. The left window shows the 'Feature_Classification_Flags' table with columns 4325, 4326, and 4327. The right window shows the same table with columns 4325, 4326, and 4327, but with different data values. Green arrows point from the 'Feature_Classification_Flags' table in the left window to the 'Feature_Classification_Flags' table in the right window. Another green arrow points from the 'Feature_Classification_Flags' table in the right window to the 'Feature_Classification_Flags' table in the left window. A third green arrow points from the 'Feature_Classification_Flags' table in the left window to the 'Feature_Classification_Flags' table in the right window. A fourth green arrow points from the 'Feature_Classification_Flags' table in the right window to the 'Feature_Classification_Flags' table in the left window. A text box with a green border contains the following text:

NetCDF JAVA API does not allow for unsigned integer data types

- The ASDC's CALIPSO Subsetter NetCDF conversion utility uses Unidata's JAVA NetCDF library (Version 4.0.49.20090825.2329) bundled with HDFView 2.9. This NetCDF library does not support unsigned integers.

The text box also contains the following metadata for 'Feature_Classification_Flags' (15300, 1):

```
16-bit unsigned integer, 225 x 5515
Number of attributes = 6
DIMENSION_LIST = 1, 15300, 12031
format = Uint_1
long_name = Feature_Classification_Flags
orig_name = Feature_Classification_Flags
units = N/A
valid_range = N/A
```

The text box also contains the following metadata for 'Feature_Classification_Flags' (7587001, 9):

```
16-bit integer, 225 x 5515
Number of attributes = 3
long_name = Feature_Classification_Flags
units = N/A
format = Uint_1
```


HDF4 to CF Compliant NetCDF Conversion Challenges

Invalid or non-compliant valid_range attribute values

- CF compliant value ranges consist of two Attribute value Array elements that are the same datatype as the attribute's variable.

Feature_Classification_Flags (15350, 2)
16-bit unsigned integer, 225 x 5515
Number of attributes = 6
DIMENSION_LIST = 1158, 12031
format = UInt_1
long_name = Feature_Classification_Flags
original_name = Feature_Classification_Flags
units = N/A
valid_range = 0

Feature_Classification_Flags (67687001, 9)
16-bit integer, 225 x 5515
Number of attributes = 3
long_name = Feature_Classification_Flags
units = N/A
format = UInt_1

HDF4 to CF Compliant NetCDF Conversion Challenges

The 'time' variable was added to address non-standard temporal parameter in the original data

- Original variable: 'Profile_UTC_Time' units = UTC - yymmdd.fxxxxxx (f is fraction of day)
- Added NetCDF CF-Compliant variable: 'time' units = days since 1970-01-01 00:00:00

Left Screenshot (Original Data):

Recent Files: C:\Users\wbaskin\Desktop\CAL_LID_L2_VFM-Prov-V2-02-2009-02-14T17-51-56ZD.hdf_Subset.nc

Table View - Feature_Classification_Flags

	4325	4326	4327
23	46107	46107	46107
24	46107	46107	46107
25	46107	46107	46107
26	46107	46107	46107
27	46107	46107	46107
28	46107	46107	46107
29	46107	46107	46107
30	46107	46107	46107
31	46107	46107	46107
32	46107	46107	46107
33	46107	46107	46107
34	46107	46107	46107
35	1	1	1

Right Screenshot (Converted Data):

Recent Files: D:\CALIPSO\data\2009-02-14\CAL_LID_L2_VFM-Prov-V2-02-2009-02-14T17-51-56ZD.hdf_Subset.nc

Table View - Feature_Classification_Flags

	4325	4326	4327
23	-19429	-19429	-19429
24	-19429	-19429	-19429
25	-19429	-19429	-19429
26	-19429	-19429	-19429
27	-19429	-19429	-19429
28	-19429	-19429	-19429
29	-19429	-19429	-19429
30	-19429	-19429	-19429
31	-19429	-19429	-19429
32	-19429	-19429	-19429
33	-19429	-19429	-19429
34	-19429	-19429	-19429
35	1	1	1

Table View - Profile_UTC_Time

	0
23	90214.75705362501
24	90214.75705223611
25	90214.75707084723
26	90214.75707945833
27	90214.75708069845
28	90214.75709668055
29	90214.75710529168
30	90214.75711389122
31	90214.75712250231
32	90214.75713111344
33	90214.75713972453
34	90214.75714833566
35	90214.75715694675
36	90214.75716555788
37	90214.75717416898
38	90214.7571827801
39	90214.7571913912
40	90214.75720000232
41	90214.75720861342
42	90214.75721722454
43	90214.75722583564
44	90214.75723444676
45	90214.75724305786
46	90214.75725166898
47	90214.75726028008
48	90214.7572688912
49	90214.75727750233
50	90214.75728611342
51	90214.75729472455
52	90214.75730332406
53	90214.75731193519
54	90214.75732054628
55	90214.75732915741
56	90214.75733776852
57	90214.75734637963
58	90214.75735499075
59	90214.75736360185
60	90214.75737221297
61	90214.75738082407
62	90214.7573894352
63	90214.75739804629
64	90214.75740665742
65	90214.75741526851

Table View - Longitude

	0
23	-58.68018
24	-58.693867
25	-58.707577
26	-58.72141
27	-58.7351
28	-58.74883
29	-58.762768
30	-58.776283
31	-58.790028
32	-58.803833
33	-58.81745
34	-58.83115
35	-58.84474
36	-58.858265
37	-58.87182
38	-58.88549
39	-58.89903
40	-58.912533
41	-58.926228
42	-58.939682
43	-58.953037
44	-58.96641
45	-58.980087
46	-58.993507
47	-59.00683
48	-59.020153
49	-59.03367
50	-59.04714
51	-59.060635
52	-59.074112
53	-59.087246
54	-59.10059
55	-59.11403
56	-59.127464
57	-59.14059
58	-59.15409
59	-59.16753
60	-59.1808
61	-59.194168
62	-59.207394
63	-59.22057
64	-59.233795
65	-59.247116

Feature_Classification_Flags (15350, 2)
16-bit unsigned integer, 225 x 5515
Number of attributes = 6
DIMENSION_LIST = 11158, 12031
format = UInt_1
long_name = Feature_Classification_Flags
orig_name = Feature_Classification_Flags
units = N/A
valid_range = N/A

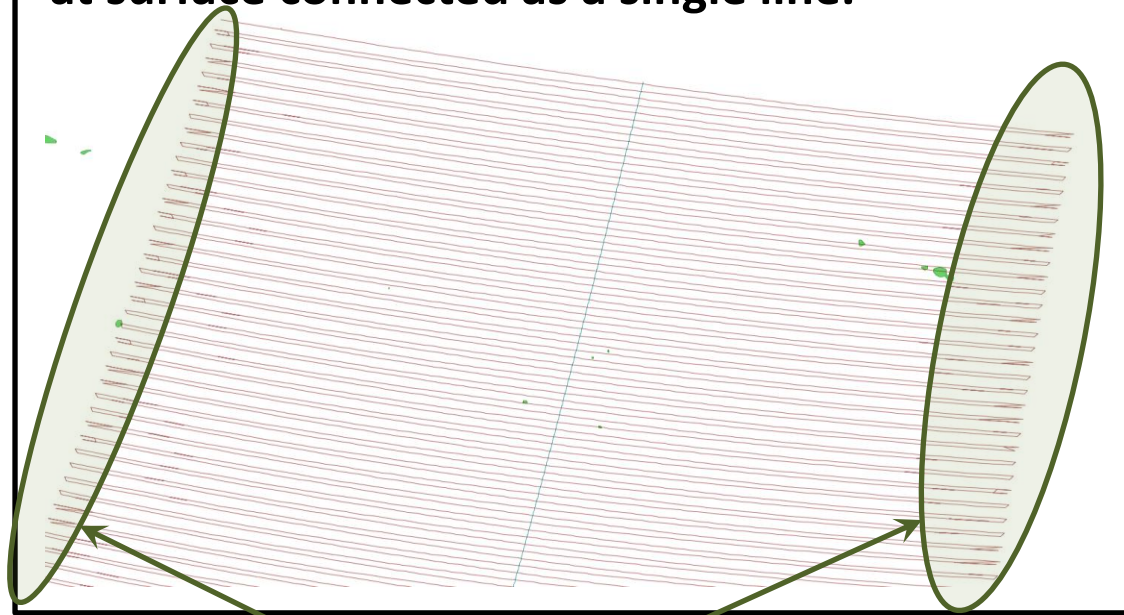
Log Info Metadata

Feature_Classification_Flags (67687001, 9)
16-bit integer, 225 x 5515
Number of attributes = 3
long_name = Feature_Classification_Flags
units = N/A
format = UInt_1

Log Info Metadata

HDF4 to CF Compliant NetCDF Conversion Challenges

Plot of CER_SSF_Aqua-FM3-MODIS Field of view at surface connected as a single line.

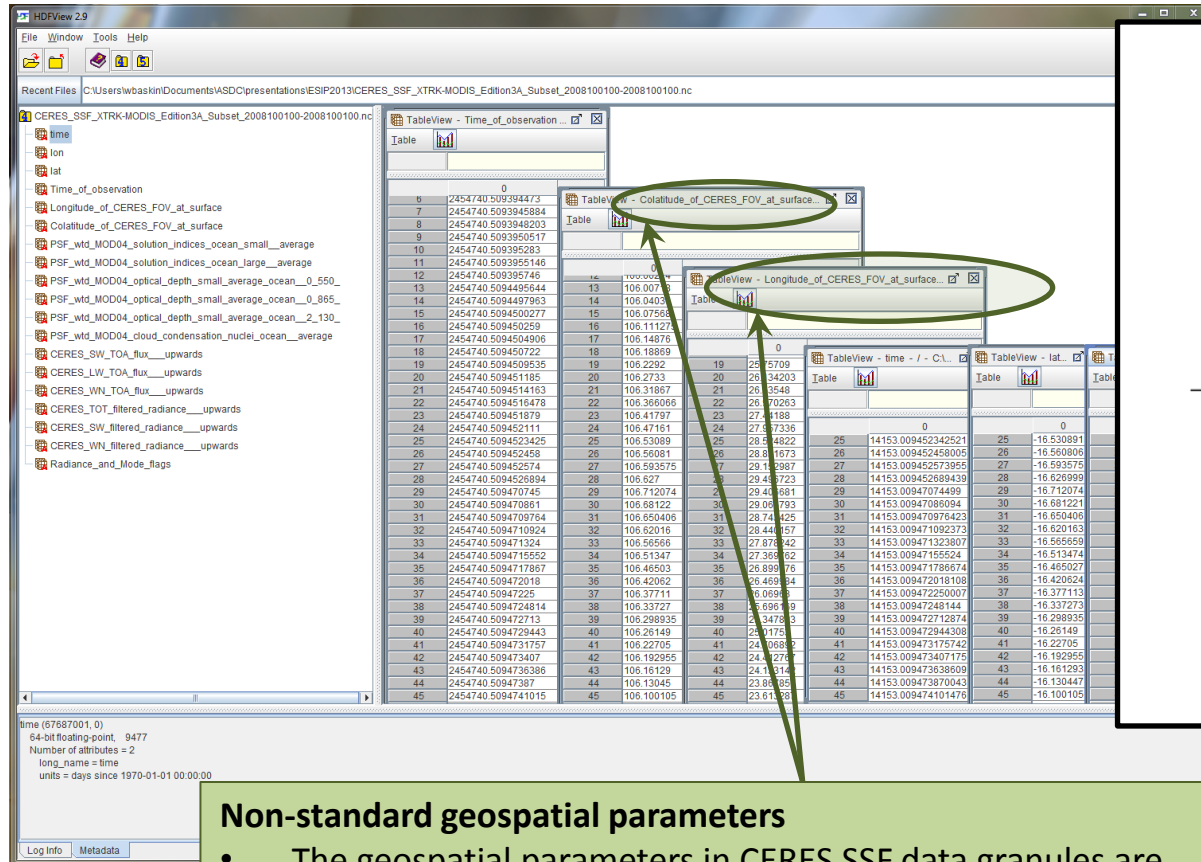


Time dimension not in temporal order

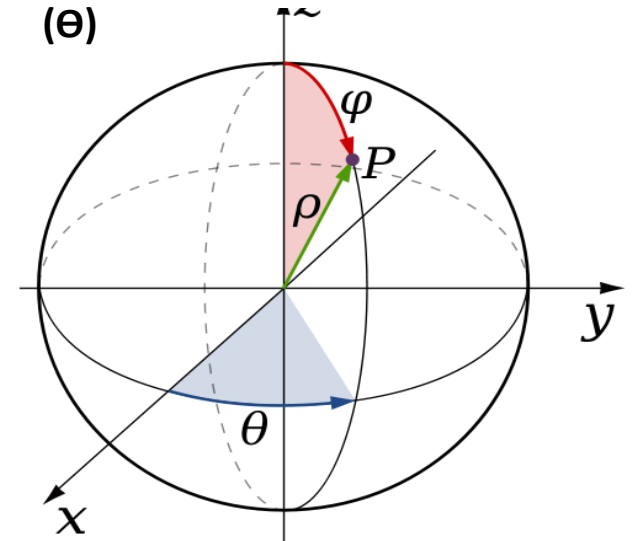
- The primary dimension in all swath-based level 2 satellite data is time. The frayed edges of the cross-track scan path illustrates that CERES SSF data is not always in temporal order. In order to be CF compliant all the data must be sorted temporally.

HDF4 to CF Compliant NetCDF Conversion Challenges

NetCDF file generated from ASDC CERES SSF Subsetter



Colatitude (ϕ) and Longitude (θ)

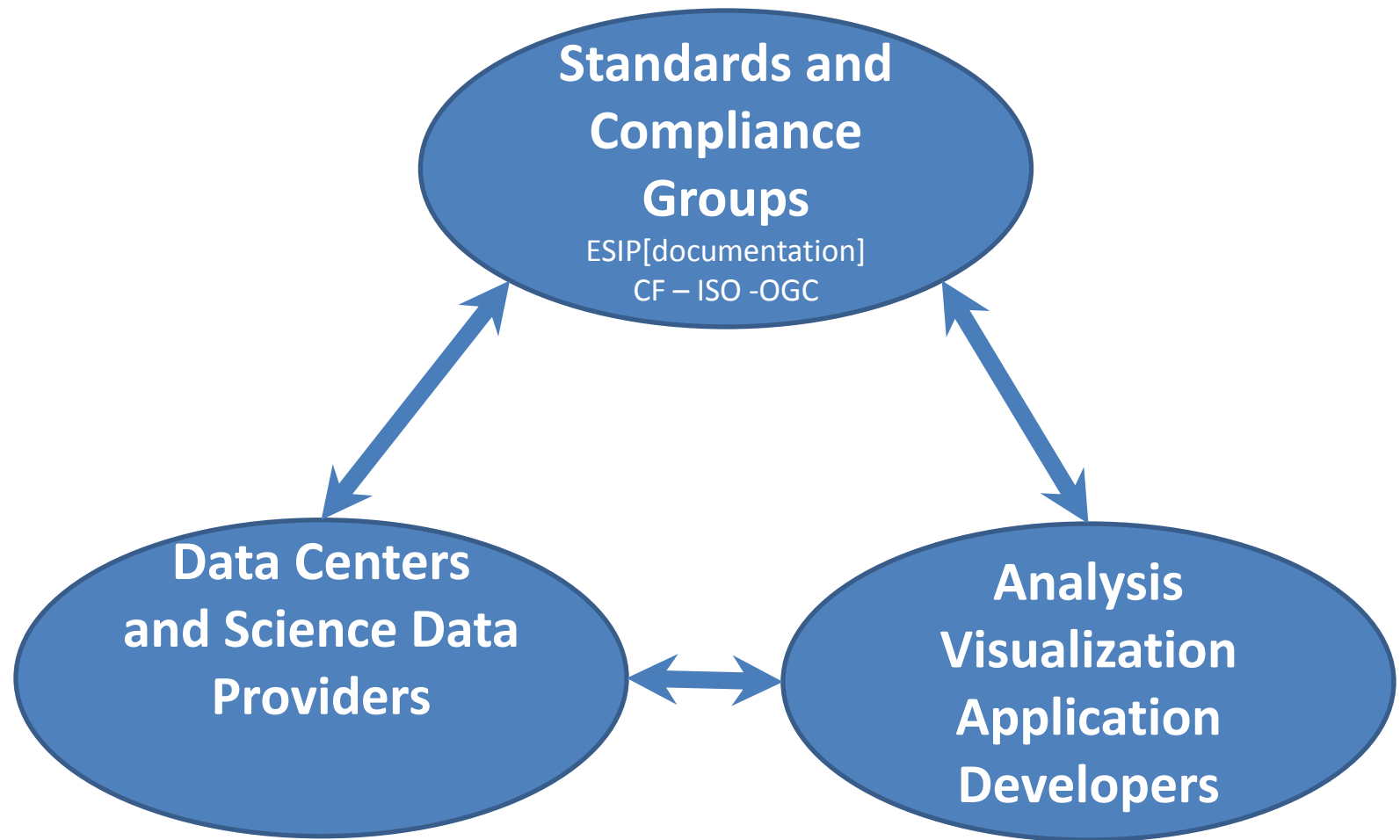


Non-standard geospatial parameters

- The geospatial parameters in CERES SSF data granules are Colatitude_of_CERES_FOV_at_surface and Longitude_of_CERES_FOV at surface. 'lon' and 'lat' parameters were added as primary CF Compliant variables and to simplify plotting in standard GIS utilities.

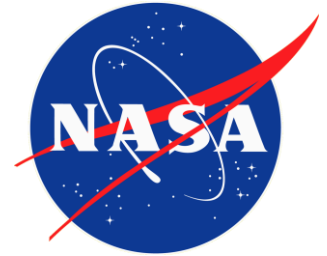
CONCLUSION

Collaborate to solve 3-way 'chicken or egg' problem





Contacts



Walt Baskin

Senior Software Engineer, ASDC (SSAI contractor)

NASA Langley Research Center

Walter.E.Baskin@nasa.gov

Office: (757) 864-8089

Jennifer Perez

Deputy Director, ASDC

NASA Langley Research Center

Jennifer.L.Perez@nasa.gov

Office: (757) 864-5329